# Robust Linearly Constrained Kalman Filter for General Mismatched Linear State-Space Models

Jordi Vilà-Valls, *Senior Member, IEEE,* Eric Chaumette, *Member, IEEE,*
François Vincent, *Member, IEEE,* and Pau Closas *Senior Member, IEEE*

***Abstract*—It is well-known that Wiener and Kalman filter (KF) like techniques are sensitive to misspecified covariances, uncertainties in the system matrices, filter initialization or unwanted system behaviors. A possible solution to robustify these estimation techniques is to impose linear constraints (LCs). In this article: i) we introduce a general class of linearly constrained KF (LCKF), where a set of non-stationary LCs can be set at every time step, ii) explore the use of such LCs to mitigate modeling errors in general mismatched linear discrete state-space models, and iii) provide the theoretical formulation to show that the gain-constrained KF is a particular instance of the proposed LCKF. Because such LCs can be taken into account in any KF generalization this sets the basis for a new robust filtering framework. An illustrative example is provided to support the discussion.**

***Index Terms*—Robust filtering, model mismatch, linearly constrained Kalman filter, mitigation, robustness, distortionless.**

## I. Introduction

The well-known Kalman filter (KF) is the recursive form of the Wiener filter (WF) for linear discrete state-space (LDSS) models. The sensitivity of the achievable mean square error (MSE) to modelling errors is a well-known intrinsic issue for any WF implementation [1]–[3]. Then the KF's achievable performance also strongly depends on the accurate knowledge of the system model, initial state and noise statistics [3]–[5]. In many applications dealing with parameter estimation [1, § 6.6], [2], linearly constrained WFs (LCWFs) have been developed in which linear constraints (LCs) are imposed to robustify WFs. Robustness is understood as the ability to achieve *close-to-optimal* performance with imperfect, incomplete or erroneous system knowledge, while minimal impact on performance under nominal conditions is caused.

In that perspective, it has been shown in [6] that adding a distortionless constraint to WF in the context of LDSS models yields to the linear minimum variance distortionless response filter (LMVDRF), a suboptimal filter in the MSE sense, but it does not depend on the prior knowledge on the initial state. Remarkably, LMVDRF exists under more general conditions than the information filter (IF) form of KF [6, §V.C], a well established solution to cope with a lack of prior initial state information [3, §6.2]. However, since LMVDRF only differs from KF in its initialization, sharing then the same recursion, it also shares the same sensitivity.

J. Vilà-Valls, E. Chaumette and F. Vincent are with ISAE-SUPAERO/University of Toulouse, Toulouse, France; P. Closas is with the ECE Dept., Northeastern University, Boston, MA, USA.

This work has been partially supported by the DGA/AID projects 2018.60.0072.00.470.75.01 and 2019.65.0068.00.470.75.01 and the NSF under Awards CNS-1815349 and ECCS-1845833.

The family of LCs for which LCWF can be computed recursively in the form of KF, leading to a linearly constrained KF (LCKF), was identified in a preliminary study [7], [8], and is discussed in Sec. III. We show that this family of LCs encompasses the gain-constrained KF (GCKF) derived for a restricted subset of real-valued LDSS models [9, (2.32)], revisited from a different perspective in Sec. IV to establish its connection to LCKF. From this new derivation we identify that once a constraint is imposed it is implicitly propagated.

We focus on the exploitation of LCs in order to robustify KF. In that context, we provide a detailed analysis of the robust LCKF capabilities against mismatches on LDSS models. This analysis relies on two particularly features of LCKF, i.e, easy to implement and fully adaptive in the context of sequential estimation. Indeed, LCKF allows optional addition of LCs that can be triggered by a preprocessing of each new measurement. This is a key feature in numerous real-life applications.

State estimation is an important component in many applications such as indoor positioning, power systems, machine learning, communication systems, vehicle control, or navigation systems to name a few [10]–[14]. So far, in those applications where the state and measurement noise statistical properties are not accurately known, it has been common practice to use $H_\infty$ filters [3, §10] [15]–[17], which do not make assumptions on the noise distribution and attempt to minimize the worst-case estimation error. Other possible ways to robustify the KF to noise mismodeling via unbiased finite impulse response (UFIR) [18], p-shift FIR [19] [20, §11] or minimum variance UFIR [21] filters, were introduced. These algorithms have the KF predictor/corrector form, may ignore initial estimation errors and the statistics of the noise, but become virtually optimal as the length of the FIR window increases, which in practice is a tuning factor. It is also worth pointing out that if the system is only affected by outliers, an efficient solution is to resort to robust statistics [22]. Other robust solutions based on moving-horizon or simultaneous state and system parameters/inputs estimation also exist [23].

LCKF provides an alternative solution to $H_\infty$ and UFIR filters to robustify KF which is worth knowing, since LCKF is able to cope with a quite general class of mismatches without need of heuristically tuning the filter parameters. However, the disadvantage of LCs is that additional degrees of freedom are used by LCKF in order to satisfy them, which increases the minimum MSE achieved. Last, LCs can be used in any existing KF generalization [3, §7]: correlated/colored noises; fading memory; state constraints; or prediction/smoothing extensions; therefore *setting the basis for a new robust filtering framework*.

## II. General Mismatched System Model

The LDSS models considered in this article are defined by the following three dynamic systems mismatched set:

$$\text{Nominal}: \begin{cases} \mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{c}_k + \mathbf{v}_k \end{cases} \tag{1a}$$

$$\text{Assumed}: \begin{cases} \mathbf{x}'_k = \widehat{\mathbf{F}}_{k-1}\mathbf{x}'_{k-1} + \widehat{\mathbf{u}}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{y}_k = \widehat{\mathbf{H}}_k\mathbf{x}'_k + \widehat{\mathbf{c}}_k + \mathbf{v}_k \end{cases} \tag{1b}$$

$$\text{True}: \begin{cases} \mathbf{x}''_k = \mathbf{F}_{k-1}\mathbf{x}''_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1} + \boldsymbol{\eta}_{k-1} \\ \mathbf{y}_k = \mathbf{H}_k\mathbf{x}''_k + \mathbf{c}_k + \mathbf{v}_k + \mathbf{j}_k \end{cases} \tag{1c}$$

which encompass several types of model mismatches. Indeed, throughout this paper, we consider the situation where a practitioner wants to estimate the state of a so-called "nominal" model (1a) which contains the information of interest. However, firstly, a perfect knowledge of the "nominal" model may not be accessible in some cases, compelling the practitioner to assume a model, the so-called "assumed" model (1b), which is believed to be as representative as possible of the "nominal" one. Secondly, some unwanted disturbances may also be present ($\boldsymbol{\eta}_{k-1}$ and/or $\mathbf{j}_k$), e.g., due to harsh environments malicious interferences or system noise statistics uncertainty, impairing the "nominal" model and leading to the actual observed model, the so-called "true" model (1c). The variables in (1a)-(1c) are: i) $\mathbf{x}_k, \mathbf{x}'_k$ and $\mathbf{x}''_k \in \mathbb{C}^{P_k}$, the nominal, assumed and true states; ii) $\mathbf{y}_k \in \mathbb{C}^{N_k}$ the observation vector at time $k$; iii) $\widehat{\mathbf{F}}_k$ and $\widehat{\mathbf{H}}_k$ the assumed model matrices. $\mathbf{F}_k$ and $\mathbf{H}_k$ are their nominal/true counterpart; iv) noise sequences $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_k\}$ are random vectors with known covariance/cross-covariance; v) $\mathbf{u}_k$, $\mathbf{c}_k$, $\widehat{\mathbf{u}}_k$ and $\widehat{\mathbf{c}}_k$ are nominal/true and assumed system inputs; and vi) $\boldsymbol{\eta}_k$ and $\mathbf{j}_k$ represent the unwanted disturbances. Standardly, the practitioner performs estimation based on the assumed knowledge, expecting that i) $\widehat{\mathbf{F}}_k$, $\widehat{\mathbf{H}}_k$, $\widehat{\mathbf{u}}_k$ and $\widehat{\mathbf{c}}_k$ are close enough to the nominal ones in (1a) to obtain reasonably good performance, and ii) there are no unwanted disturbances ($\boldsymbol{\eta}_{k-1} = \mathbf{0}$ and $\mathbf{j}_k = \mathbf{0}$), that is, (1a)=(1c). However, in practice other situations may arise, and we can distinguish three general scenarios: 1) {Assumed = True}, the ideal case with no mismatch and no unwanted disturbances, 2) {True = Nominal}, no unwanted disturbances but possible mismatch, and 3) {Assumed = Nominal}, no mismatch but possible unwanted disturbances. In any of these situations, *the ultimate goal is to estimate the nominal state* $\mathbf{x}_k$ *(1a), which is the information of interest, based on the (true) measurements* $\mathbf{y}_k$ *(1c) and our knowledge of the model dynamics (1b).*

## III. Linearly Constrained Wiener/Kalman Filter

We consider first the ideal case where {Assumed = True} ((1b)=(1c) ⇒ (1b)=(1a)) with state/measurement equations,

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{2a}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{2b}$$

with $k \geq 1$, and $\mathbf{F}_k$, $\mathbf{H}_k$ known model matrices. Noise sequences, as well as $\mathbf{x}_0$, are random vectors with known covariance and cross-covariance. Then the linear minimum MSE (LMMSE) estimator of $\mathbf{x}_k$ ($k \geq 2$) is the WF[1]

$$\widehat{\mathbf{x}}^b_{k|k} = \mathbb{K}^b_k\overline{\mathbf{y}}_k,\ \mathbb{K}^b_k = \arg\min_{\mathbb{K}_k}\{\mathbf{P}_{k|k}(\mathbb{K}_k)\} = \mathbf{C}_{\mathbf{x}_k,\overline{\mathbf{y}}_k}\mathbf{C}^{-1}_{\overline{\mathbf{y}}_k},$$
$$\mathbf{P}_{k|k}(\mathbb{K}_k) = \mathbb{E}[(\mathbb{K}_k\overline{\mathbf{y}}_k - \mathbf{x}_k)(\cdot)^H], \tag{3}$$

where $\overline{\mathbf{y}}^\top_k = (\mathbf{y}^\top_1, \ldots, \mathbf{y}^\top_k) \in \mathbb{C}^{\mathcal{N}_k}$, $\mathcal{N}_k = \sum_{l=1}^k N_l$. Since Kalman's work [3], it is known that, if $\{\mathbf{w}_k, \mathbf{v}_k, \mathbf{x}_0\}$ verify certain uncorrelation conditions, lately extended in [6],

$$\mathbf{C}_{\mathbf{w}_{k-1},\overline{\mathbf{y}}_{k-1}} = \mathbf{0},\ \mathbf{C}_{\mathbf{v}_k,\overline{\mathbf{y}}_{k-1}} = \mathbf{0},\ \forall k \geq 2, \tag{4}$$

then $\widehat{\mathbf{x}}^b_{k|k}$ (3) admits a convenient recursive form ($k \geq 2$),

$$\widehat{\mathbf{x}}^b_{k|k} = \left(\mathbf{I} - \mathbf{K}^b_k\mathbf{H}_k\right)\mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} + \mathbf{K}^b_k\mathbf{y}_k \tag{5}$$

so-called KF estimate of $\mathbf{x}_k$ [3], where the gain $\mathbf{K}^b_k \in \mathbb{C}^{P_k \times N_k}$ verifies $\mathbf{K}^b_k = \arg\min_{\mathbf{K}_k}\{\mathbf{P}^J_{k|k}(\mathbf{K}_k)\}$, with $\mathbf{P}^J_{k|k}(\mathbf{K}_k)$ the general form of Joseph covariance update equation,

$$\begin{aligned}\mathbf{P}^J_{k|k}(\mathbf{K}_k) &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^b_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^H \\ &+ \mathbf{K}_k\mathbf{C}_{\mathbf{v}_k}\mathbf{K}^H_k - (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k}\mathbf{K}^H_k \\ &- \mathbf{K}_k\mathbf{C}^H_{\mathbf{x}_k,\mathbf{v}_k}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^H.\end{aligned} \tag{6}$$

Thus $\mathbf{K}^b_k$ is computed in its most general form as [6],

$$\begin{aligned}\mathbf{P}^b_{k|k-1} &= \mathbf{F}_{k-1}\mathbf{P}^b_{k-1|k-1}\mathbf{F}^H_{k-1} + \mathbf{C}_{\mathbf{w}_{k-1}} \\ &+ \mathbf{F}_{k-1}\mathbf{C}^H_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}\mathbf{F}^H_{k-1}\end{aligned} \tag{7a}$$

$$\mathbf{S}^b_{k|k-1} = \mathbf{H}_k\mathbf{P}^b_{k|k-1}\mathbf{H}^H_k + \mathbf{C}_{\mathbf{v}_k} + \mathbf{H}_k\mathbf{C}^H_{\mathbf{v}_k,\mathbf{x}_k} + \mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}\mathbf{H}^H_k$$

$$\mathbf{K}^b_k = (\mathbf{P}^b_{k|k-1}\mathbf{H}^H_k + \mathbf{C}^H_{\mathbf{v}_k,\mathbf{x}_k})(\mathbf{S}^b_{k|k-1})^{-1} \tag{7b}$$

$$\mathbf{P}^b_{k|k} = (\mathbf{I} - \mathbf{K}^b_k\mathbf{H}_k)\mathbf{P}^b_{k|k-1} - \mathbf{K}^b_k\mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}. \tag{7c}$$

The above recursion (5)-(7c) is also valid for $k = 1$ provided that $\mathbf{P}^b_{0|0} = \mathbf{C}_{\mathbf{x}_0}$ and $\widehat{\mathbf{x}}^b_{0|0} = \mathbf{0}$ [6]. Notice that the "standard LDSS model" in monographs [24, §9.1] [3, §7.1] satisfies

$$\begin{aligned}&\mathbf{C}_{\mathbf{x}_0,\mathbf{w}_k} = \mathbf{0},\ \mathbf{C}_{\mathbf{x}_0,\mathbf{v}_k} = \mathbf{0},\ \mathbf{C}_{\mathbf{w}_l,\mathbf{w}_k} = \mathbf{C}_{\mathbf{w}_k}\delta^l_k, \\ &\mathbf{C}_{\mathbf{v}_l,\mathbf{v}_k} = \mathbf{C}_{\mathbf{v}_k}\delta^l_k,\ \mathbf{C}_{\mathbf{w}_l,\mathbf{v}_k} = \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}\delta^{l+1}_k,\end{aligned} \tag{8}$$

which has long been regarded as leading to the general form of KF including correlated process and measurement noise, and is in fact a special case of (4).

### A. Linearly Constrained WF/KF Formulation

We assume that the WF estimate of $\mathbf{x}_k$ (3) exists which implies that $\mathbf{C}_{\overline{\mathbf{y}}_k}$ is invertible. Introducing a set of LCs, i.e., $\mathbb{K}_k\overline{\boldsymbol{\Lambda}}_k = \mathbf{T}_k$, into model (3) yields a LCWF [2],

$$\widehat{\mathbf{x}}^b_{k|k} = \mathbb{L}^b_k\overline{\mathbf{y}}_k,\ \mathbb{L}^b_k = \arg\min_{\mathbb{L}_k}\{\mathbf{P}_{k|k}(\mathbb{L}_k)\} \text{ s.t. } \mathbb{L}_k\overline{\boldsymbol{\Lambda}}_k = \mathbf{T}_k,$$
$$\mathbb{L}^b_k = \mathbb{K}^b_k + (\mathbf{T}_k - \mathbb{K}^b_k\overline{\boldsymbol{\Lambda}}_k)(\overline{\boldsymbol{\Lambda}}^H_k\mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\overline{\boldsymbol{\Lambda}}_k)^{-1}\overline{\boldsymbol{\Lambda}}^H_k\mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}. \tag{9}$$

Let us assume the following block matrix decomposition: $\mathbb{L}_k = [\mathbb{J}_{k-1}\ \mathbf{L}_k]$ where $\mathbb{J}_{k-1} \in \mathbb{C}^{P_k \times \mathcal{N}_{k-1}}$ and $\mathbf{L}_k \in \mathbb{C}^{P_k \times N_k}$ leading to $\mathbb{L}_k\overline{\mathbf{y}}_k = \mathbb{J}_{k-1}\overline{\mathbf{y}}_{k-1} + \mathbf{L}_k\mathbf{y}_k$. It appears [7] that the

[1] In Sec. III and IV, for the sake of legibility and without loss of generality, we assume that $\mathbb{E}[\mathbf{x}_0] = \mathbf{0}$ and no system inputs i.e., $\mathbf{u}_{k-1} = \mathbf{0}$ and $\mathbf{c}_k = \mathbf{0}$, in the KF/LCKF derivation. The estimate of $\mathbf{x}_k$ based on measurements up to time $k$ is denoted $\widehat{\mathbf{x}}_{k|k} \triangleq \widehat{\mathbf{x}}_{k|k}(\mathbf{y}_1, \ldots, \mathbf{y}_k)$. The superscript $(.)^b$ denotes that the considered value is the "best" one according to a given criterion.

subset of LCs $\mathbb{L}_k\overline{\mathbf{\Lambda}}_k = \mathbf{T}_k$ leading to a recursion similar to (5) for the LCWF estimate of $\mathbf{x}_k$ (9), that is,

$$\widehat{\mathbf{x}}^b_{k|k} = \left(\mathbf{I} - \mathbf{L}^b_k\mathbf{H}_k\right)\mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} + \mathbf{L}^b_k\mathbf{y}_k,\ k \geq 1, \quad (10)$$

consists of the following three types of LCs:

$$\mathcal{C}^1_k : [\mathbb{J}_{k-1}\ \mathbf{L}_k]\begin{bmatrix}\mathbf{0}\\ \mathbf{\Delta}_k\end{bmatrix} = \mathbf{T}_k \quad (11a)$$

$$\mathcal{C}^2_k : [\mathbb{J}_{k-1}\ \mathbf{L}_k]\begin{bmatrix}\overline{\mathbf{\Lambda}}_{k-1}\\ \mathbf{H}_k\mathbf{F}_{k-1}\mathbf{T}_{k-1}\end{bmatrix} = \mathbf{F}_{k-1}\mathbf{T}_{k-1} \quad (11b)$$

$$\mathcal{C}^3_k : [\mathbb{J}_{k-1}\ \mathbf{L}_k]\begin{bmatrix}\overline{\mathbf{\Lambda}}_{k-1} & \mathbf{0}\\ \mathbf{H}_k\mathbf{F}_{k-1}\mathbf{T}_{k-1} & \mathbf{\Delta}_k\end{bmatrix} = [\mathbf{F}_{k-1}\mathbf{T}_{k-1}\ \mathbf{T}_k] \quad (11c)$$

where:

- *$\mathcal{C}^1_k$ is dedicated to introduce a first subset of LCs at $k$,*
- *$\mathcal{C}^2_k$ corresponds to the implicit propagation at $k$ via recursion (10) of the LCs already set from $1$ to $k-1$,*
- *$\mathcal{C}^3_k$ combines $\mathcal{C}^2_k$ and $\mathcal{C}^1_k$, propagation of previously set LCs before $k$ and addition of a new subset of LCs at $k$.*

Under $\mathcal{C}^2_k$, the gain is given by $\mathbf{L}^b_k = \arg\min_{\mathbf{L}_k}\{\mathbf{P}^J_{k|k}(\mathbf{L}_k)\}$, and can be computed from the "unconstrained" KF recursion (7a)-(7c). Under $\mathcal{C}^1_k$ and $\mathcal{C}^3_k$, $\mathbf{L}^b_k$ is the solution of

$$\mathbf{L}^b_k = \arg\min_{\mathbf{L}_k}\{\mathbf{P}^J_{k|k}(\mathbf{L}_k)\} \text{ s.t. } \mathbf{L}_k\mathbf{\Delta}_k = \mathbf{T}_k, \quad (12)$$

which is given by ($\mathbf{P}^b_{k|k-1}$, $\mathbf{S}^b_{k|k-1}$ and $\mathbf{K}_k$ as in (7a)-(7b))

$$\mathbf{\Gamma}_k = \mathbf{T}_k - \mathbf{K}_k\mathbf{\Delta}_k, \quad \mathbf{\Psi}_k = \mathbf{\Delta}^H_k(\mathbf{S}^b_{k|k-1})^{-1}\mathbf{\Delta}_k \quad (13a)$$

$$\mathbf{L}^b_k = \mathbf{K}_k + \mathbf{\Gamma}_k\mathbf{\Psi}^{-1}_k\mathbf{\Delta}^H_k(\mathbf{S}^b_{k|k-1})^{-1} \quad (13b)$$

$$\mathbf{P}^b_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^b_{k|k-1} - \mathbf{K}_k\mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k} + \mathbf{\Gamma}_k\mathbf{\Psi}^{-1}_k\mathbf{\Gamma}^H_k \quad (13c)$$

Interestingly, the recursive formulation of LCWF for LDSS models introduced is fully adaptive in the context of sequential estimation as it allows at each new observation to incorporate or not new LCs. Since KF is the recursive form of WF for LDSS models, it makes sense to denote LCKF the recursive form of LCWF for such LDSS models.

### B. Validity of the Predictor/Corrector Form

The general LCKF can be revisited by expressing the KF recursion in the so-called "predictor/corrector" form, that is, as a function of the a priori estimates of $\mathbf{x}_k$ and $\mathbf{y}_k$, for $k \geq 1$,

$$\widehat{\mathbf{x}}^b_{k|k} = \widehat{\mathbf{x}}^b_{k|k-1} + \mathbf{K}^b_k\boldsymbol{\varepsilon}^b_{k|k-1} \quad (14a)$$

$$\boldsymbol{\varepsilon}^b_{k|k-1} = \mathbf{y}_k - \widehat{\mathbf{y}}^b_{k|k-1} \quad (14b)$$

$$\widehat{\mathbf{y}}^b_{k|k-1} = \mathbf{H}_k\widehat{\mathbf{x}}^b_{k|k-1},\ \widehat{\mathbf{x}}^b_{k|k-1} = \mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1}, \quad (14c)$$

where $\mathbf{P}^b_{k|k-1} = E[(\widehat{\mathbf{x}}^b_{k|k-1} - \mathbf{x}_k)(\cdot)^H]$. Then, $\boldsymbol{\varepsilon}^b_{k|k-1}$ is called the innovations' vector and $\mathbf{S}^b_{k|k-1} = E[\boldsymbol{\varepsilon}^b_{k|k-1}\boldsymbol{\varepsilon}^{bH}_{k|k-1}]$. Interestingly enough, this formalism is still valid for LCKF. Indeed, if LCs are set between time 1 and time $k$, LCKF of $\mathbf{x}_k$ is the solution to

$$\widehat{\mathbf{x}}^b_{k|k} = \mathbb{L}^b_k\overline{\mathbf{y}}_k,\ \mathbb{L}^b_k = \arg\min_{\mathbb{L}_k}\{\mathbf{P}_{k|k}(\mathbb{L}_k)\} \text{ s.t. } \mathbb{L}_k\overline{\mathbf{\Lambda}}_k = \mathbf{T}_k.$$

Then, leveraging (9), it is easy to check that

$$\mathbf{F}_k\mathbb{L}^b_k = \arg\min_{\mathbb{M}_k}\{\mathbb{E}[\widehat{\boldsymbol{\delta}}(\mathbb{M}_k)\widehat{\boldsymbol{\delta}}^H(\mathbb{M}_k)]\} \text{ s.t.}$$
$$\mathbb{M}_k\overline{\mathbf{\Lambda}}_k = \mathbf{F}_k\mathbf{T}_k, \quad \widehat{\boldsymbol{\delta}}(\mathbb{M}_k) = \mathbb{M}_k\overline{\mathbf{y}}_k - \mathbf{F}_k\mathbf{x}_k.$$

Moreover, under (4) we have that

$$\mathbf{P}_{k+1|k}(\mathbb{M}_k) = \mathbb{E}[(\mathbb{M}_k\overline{\mathbf{y}}_k - \mathbf{x}_{k+1})(\cdot)^H]$$
$$= \mathbf{C}_{\mathbf{w}_k} + \mathbf{F}_k\mathbf{C}_{\mathbf{x}_k,\mathbf{w}_k} + \mathbf{C}_{\mathbf{w}_k,\mathbf{x}_k}\mathbf{F}^H_k + \mathbb{E}[\widehat{\boldsymbol{\delta}}(\mathbb{M}_k)\widehat{\boldsymbol{\delta}}^H(\mathbb{M}_k)],$$

and $\mathbf{F}_k\mathbb{L}^b_k = \arg\min_{\mathbb{M}_k}\{\mathbf{P}_{k+1|k}(\mathbb{M}_k)\}$ s.t. $\mathbb{M}_k\overline{\mathbf{\Lambda}}_k = \mathbf{F}_k\mathbf{T}_k$, which means that $\widehat{\mathbf{x}}^b_{k+1|k} = \mathbf{F}_k\widehat{\mathbf{x}}^b_{k|k} = \mathbf{F}_k\mathbb{L}^b_k\overline{\mathbf{y}}_k$ has become a constrained a priori estimate of $\mathbf{x}_k$. Likewise, it follows that

$$\mathbf{H}_{k+1}\mathbf{F}_k\mathbb{L}^b_k = \arg\min_{\mathbb{M}_k}\{\mathbb{E}[\widehat{\boldsymbol{\zeta}}(\mathbb{M}_k)\widehat{\boldsymbol{\zeta}}(\mathbb{M}_k)^H]\} \text{ s.t.}$$
$$\mathbb{M}_k\overline{\mathbf{\Lambda}}_k = \mathbf{H}_k\mathbf{F}_k\mathbf{T}_k, \quad \widehat{\boldsymbol{\zeta}}(\mathbb{M}_k) = \mathbb{M}_k\overline{\mathbf{y}}_k - \mathbf{y}_{k+1},$$

which means that $\widehat{\mathbf{y}}^b_{k+1|k} = \mathbf{H}_{k+1}\mathbf{F}_k\widehat{\mathbf{x}}^b_{k|k} = \mathbf{H}_{k+1}\mathbf{F}_k\mathbb{L}^b_k\overline{\mathbf{y}}_k$ has become a constrained a priori estimate of $\mathbf{y}_k$, leading to a constrained innovations' vector $\boldsymbol{\varepsilon}^b_{k+1|k} = \mathbf{y}_{k+1} - \widehat{\mathbf{y}}^b_{k+1|k}$.

In conclusion, *once a subset of LCs has been introduced by $\mathcal{C}^1_k$, all the subsequent filter estimates, a priori estimates and innovations' vector are linearly constrained, even if no new subset of LCs are explicitly introduced via $\mathcal{C}^3_k$*. Notice that this is a noteworthy property of LCKFs which cannot be highlighted with the gain-constrained approach in [9].

## IV. LCKF versus Gain-constrained KF

### A. Revisiting the Gain-constrained KF

Let us consider LCKF defined by the sequence of LCs $\{\mathcal{C}^1_1, \mathcal{C}^3_2, \ldots, \mathcal{C}^3_{k-1}\}$ such that (11a)-(11c),

$$\mathbf{\Delta}_k = \mathbf{E}_k, \quad \mathbf{T}_k = \mathbf{D}^R_k\mathbf{F}_k + (\mathbf{I} - \mathbf{\Pi}_k)\left(\mathbf{K}_k\mathbf{E}_k - \mathbf{D}^R_k\mathbf{F}_k\right),$$

where $\mathbf{K}_k$ is the classical Kalman gain (7b), $\mathbf{\Pi}_k = \mathbf{\Xi}^{-1}_k\mathbf{D}^H_k\left(\mathbf{D}_k\mathbf{\Xi}^{-1}_k\mathbf{D}^H_k\right)^{-1}\mathbf{D}_k$ is an oblique projector on $\mathbf{\Xi}^{-1}_k\mathbf{D}^H_k$, $\mathbf{\Xi}_k \in \mathbb{C}^{P_k \times P_k}$ is a Hermitian positive definite weighting matrix, $\mathbf{D}_k \in \mathbb{C}^{Q_k \times P_k}$, $Q_k \leq P_k$, is a right invertible matrix whose right inverse is $\mathbf{D}^R_k = \mathbf{D}^H_k(\mathbf{D}_k\mathbf{D}^H_k)^{-1}$, $\mathbf{E}_k \in \mathbb{C}^{N_k \times R_k}$, $R_k \leq N_k$, is a left invertible matrix whose left inverse is $\mathbf{E}^L_k = (\mathbf{E}^H_k\mathbf{E}_k)^{-1}\mathbf{E}^H_k$, and $\mathbf{F}_k \in \mathbb{C}^{Q_k \times R_k}$. Then, from (13b):

$$\mathbf{L}^b_k = \mathbf{K}_k + (\mathbf{T}_k - \mathbf{K}_k\mathbf{E}_k)\mathbf{\Psi}^{-1}_k\mathbf{E}^H_k(\mathbf{S}^b_{k|k-1})^{-1}$$
$$= \mathbf{K}_k + (\mathbf{T}_k - \mathbf{K}_k\mathbf{E}_k)\mathbf{E}^L_k\mathbf{\Omega}_k$$
$$\mathbf{\Omega}_k = \mathbf{E}_k(\mathbf{E}^H_k(\mathbf{S}^b_{k|k-1})^{-1}\mathbf{E}_k)^{-1}\mathbf{E}^H_k(\mathbf{S}^b_{k|k-1})^{-1}$$

where $\mathbf{\Omega}_k$ is an oblique projector on $\mathbf{E}_k$, that is,

$$\mathbf{L}^b_k = \mathbf{K}_k - \mathbf{\Pi}_k\left(\mathbf{K}_k - \mathbf{D}^R_k\mathbf{F}_k\mathbf{E}^L_k\right)\mathbf{\Omega}_k,\ \mathbf{D}_k\mathbf{L}^b_k\mathbf{E}_k = \mathbf{F}_k,$$

which is also the solution of [9, (2.6)-(2.7)]:

$$\mathbf{L}^b_k = \arg\min_{\mathbf{L}_k}\{tr(\mathbf{P}^J_{k|k}(\mathbf{L}_k)\mathbf{\Xi}_k)\} \text{ s.t. } \mathbf{D}_k\mathbf{L}_k\mathbf{E}_k = \mathbf{F}_k, \quad (15)$$

so-called GCKF [9, (2.32)], derived for real-valued LDSS models in the restricted case where $P_k = P$, $N_k = N$, $Q_k = Q$, $R_k = R$ and the "minimal" uncorrelation conditions [3, §7.1], which is the restricted case of the "standard LDSS

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3132890, IEEE Transactions on Automatic Control



model" (8) where $\mathbf{C}_{\mathbf{w}_l,\mathbf{v}_k} = \mathbf{0}$. Hence, GCKF (15) is actually a particular case of LCKF. This observation allows to generalize the existence of GCKF to the complex-valued LDSS models under more general uncorrelation conditions (4).

### B. An Insightful Understanding of KF Behavior under LCs

In comparison with GCKF released in [9], the proposed approach, that is, the identification of LCs (11a)-(11c) for which LCWF can be formulated in a recursive LCKF form, is more general and yields to additional insights: i) from [9] it is not obvious to link GCKF with LCWF, since the formulation of LCs as in (15) cannot be easily linked to LCs as in (11c), ii) the key role of (11b) within the "unconstrained" KF recursion, in the propagation of previously set LCs, cannot be highlighted by the unconstrained solution of (15), $\mathbf{L}_k^b = \arg\min_{\mathbf{L}_k}\{(\widehat{\mathbf{x}}_{k|k} - \mathbf{x}_k)^H \mathbf{\Xi}_k (\widehat{\mathbf{x}}_{k|k} - \mathbf{x}_k)\}$. As a consequence, neither the main result of Section III-B nor the connection of the LMVDRF with the LCKF, and hence with the GCKF, can be established with the original GCKF approach. From a broader perspective, from [9] the GCKF cannot be linked to the recursive linearly constrained minimum variance filter (LCMVF) recently proposed in [8]. Indeed, the recursive LCMVF is the special case of LCKF incorporating a distortionless constraint [8], which simply amounts to consider at time $k = 1$ LCs of the form $\mathbf{L}_1\mathbf{\Lambda}_1 = \mathbf{\Gamma}_1$, $\{\mathbf{\Lambda}_1 = \mathbf{H}_1,\ \mathbf{\Gamma}_1 = \mathbf{I}\}$ or $\{\mathbf{\Lambda}_1 = [\mathbf{H}_1\ \mathbf{\Omega}_1],\ \mathbf{\Gamma}_1 = [\mathbf{I}\ \mathbf{\Upsilon}_1]\}$, since the distortionless property propagates through (11b).

Once the connection between LCMVF and LCKF has been made, it opens access to the abundant literature on LCMVF in the deterministic framework [1, § 6.7], where LCMVF has been developed to robustify MVDRF against various types of measurement mismatch. For instance, in array processing LCMVF is used to cope with i) array perturbation and/or direction of arrival mismatch leading to a misspecified $\mathbf{H}_k$, and ii) jamming and/or inaccurate estimation of thermal noise, leading to a misspecified $\mathbf{C}_{\mathbf{v}_k}$.

Without the connection between GCKF and LCMVF, authors of [9] have mainly investigated the ability of GCKF to reformulate the solution of known state estimation problems, such as: i) enforcing a linear equality constraint in the state vector, ii) enforcing unbiased estimation for systems with unknown inputs, and iii) simplification of the estimator structure for large-scale systems (yielding the spatially constrained KF).

## V. Mitigation of Modelling Errors in General Mismatched LDSS Models

The previous recursive WF (KF) and LCKF are obtained in the ideal case (no mismatch, no unwanted disturbances) where the LDSS models are of the form (2a)-(2b), that is, the three system models in Section II, (1a)-(1c) coincide. In the general case, $\mathbb{E}[\mathbf{x}_0] \neq \mathbf{0}$ or $\mathbf{u}_{k-1} \neq \mathbf{0}$ or $\mathbf{c}_k \neq \mathbf{0}$; thus one has to resort to the affine formulation of the WF which takes into account known mean values and simply yields $\widehat{\mathbf{x}}_{0|0}^b = \mathbb{E}[\mathbf{x}_0]$, $\widehat{\mathbf{x}}_{k|k-1}^b = \mathbf{F}_{k-1}\widehat{\mathbf{x}}_{k-1|k-1}^b + \mathbf{u}_{k-1}$, $\widehat{\mathbf{y}}_{k|k-1}^b = \mathbf{H}_k\widehat{\mathbf{x}}_{k|k-1}^b + \mathbf{c}_k$, whatever (unconstrained) KF or LCKF is considered. Therefore, the estimation error in the ideal case is always

$$\widehat{\mathbf{x}}_{k|k}(\mathbf{L}_k) - \mathbf{x}_k = \widehat{\mathbf{x}}_{k|k-1}^b - \mathbf{x}_k + \mathbf{L}_k(\mathbf{y}_k - \widehat{\mathbf{y}}_{k|k-1}^b) = \mathbf{L}_k\mathbf{v}_k + (\mathbf{I} - \mathbf{L}_k\mathbf{H}_k)(\mathbf{F}_{k-1}((\widehat{\mathbf{x}}_{k-1|k-1}^b - \mathbf{x}_{k-1}) - \mathbf{w}_{k-1}), \quad (16)$$

leading to the general form of Joseph covariance update equation $\mathbf{P}_{k|k}^J(\mathbf{L}_k)$ (6), which can be minimized in the unconstrained case (7a)-(7c) or in its constrained counterpart (13a)-(13c), computed with the LDSS model (2a)-(2b). In real-life applications, the system models in (1a)-(1c) may not coincide, and therefore the goal is to estimate the nominal state $\mathbf{x}_k$ (1a) based on the (true) measurements $\mathbf{y}_k$ (1c) and our model knowledge (1b). Let us denote

$$\begin{cases} \mathbf{F}_{k-1} = \widehat{\mathbf{F}}_{k-1} + d\mathbf{F}_{k-1}\ ;\ \mathbf{H}_k = \widehat{\mathbf{H}}_k + d\mathbf{H}_k \\ \mathbf{u}_{k-1} = \widehat{\mathbf{u}}_{k-1} + d\mathbf{u}_{k-1}\ ;\ \mathbf{c}_k = \widehat{\mathbf{c}}_k + d\mathbf{c}_k \end{cases},$$

and $d\mathbb{F}_{k-1} = [d\mathbf{F}_{k-1}\ d\mathbf{u}_{k-1}]$, $d\mathbb{H}_k = [d\mathbf{H}_k\ d\mathbf{c}_k]$, which gather the possible modelling errors on state, respectively measurement, matrices and/or input values. Recall that we are not looking to follow the approach in [9] to constrain the state or measurements, but rather exploit the use of LCs to robustify KF under model mismatch, thus leading to a new robust LCKF framework for the mismatched set in (1a)-(1c).

### A. Impact of Modelling Errors in Mismatched LDSS Models

At time $k \geq 2$, provided that LCs (11a)-(11c) are considered, any LCKF of $\mathbf{x}_k$ is given by the Kalman-like recursion

$$\widehat{\mathbf{x}}_{k|k}(\mathbf{L}_k) = (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)\widehat{\mathbf{x}}_{k|k-1}^b + \mathbf{L}_k(\mathbf{y}_k - \widehat{\mathbf{c}}_k) \quad (17)$$
$$\widehat{\mathbf{x}}_{k|k-1}^b = \widehat{\mathbf{F}}_{k-1}\widehat{\mathbf{x}}_{k-1|k-1}^b + \widehat{\mathbf{u}}_{k-1},$$

where again $\mathbf{L}_k \triangleq \mathbf{L}_k^b$ is obtained from the unconstrained (7a)-(7c) or constrained (13a)-(13c) recursions, but in this case computed with the assumed LDSS model (1b). Incorporating (1a)-(1c) in (17) yields the "mismatched" form of (16)

$$\widehat{\mathbf{x}}_{k|k}(\mathbf{L}_k) - \mathbf{x}_k = \mathbf{L}_k\mathbf{v}_k + \quad (18a)$$
$$(\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)(\widehat{\mathbf{F}}_{k-1}(\widehat{\mathbf{x}}_{k-1|k-1}^b - \mathbf{x}_{k-1}) - \mathbf{w}_{k-1}) + \boldsymbol{\varepsilon}_k(\mathbf{L}_k),$$

where $\boldsymbol{\varepsilon}_k(\mathbf{L}_k)$ in its most general form is given by

$$\begin{aligned} \boldsymbol{\varepsilon}_k(\mathbf{L}_k) = {} & \mathbf{L}_k\widehat{\mathbf{H}}_k(\widehat{\mathbf{F}}_{k-1} + d\mathbf{F}_{k-1})\left(\mathbf{x}_{k-1}'' - \mathbf{x}_{k-1}\right) \\ & - (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)\left(d\mathbf{F}_{k-1}\mathbf{x}_{k-1} + d\mathbf{u}_{k-1}\right) \\ & + \mathbf{L}_k d\mathbf{H}_k\left(\mathbf{w}_{k-1} + \widehat{\mathbf{u}}_{k-1} + d\mathbf{u}_{k-1}\right) \\ & + \mathbf{L}_k d\mathbf{H}_k((\widehat{\mathbf{F}}_{k-1} + d\mathbf{F}_{k-1})\mathbf{x}_{k-1}'' + \boldsymbol{\eta}_{k-1}) \\ & + \mathbf{L}_k(\widehat{\mathbf{H}}_k\boldsymbol{\eta}_{k-1} + \mathbf{j}_k + d\mathbf{c}_k). \end{aligned} \quad (18b)$$

Then, if the subset of gain matrices

$$\mathcal{L}_k = \{\mathbf{L} \in \mathbb{C}^{P_k \times N_k} \mid \boldsymbol{\varepsilon}_k(\mathbf{L}) = \mathbf{0}\}, \quad (19)$$

is non empty, then for any $\mathbf{L}_k \in \mathcal{L}_k$ (18a) reduces to

$$\widehat{\mathbf{x}}_{k|k}(\mathbf{L}_k) - \mathbf{x}_k = \mathbf{L}_k\mathbf{v}_k + (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)(\widehat{\mathbf{F}}_{k-1}(\widehat{\mathbf{x}}_{k-1|k-1}^b - \mathbf{x}_{k-1}) - \mathbf{w}_{k-1}). \quad (20)$$

Without any additional assumptions, the best $\mathbf{L}_k \in \mathcal{L}_k$ in the MSE sense is the one that verifies

$$\mathbf{L}_k^b = \arg\min_{\mathbf{L}_k}\{\mathbf{P}_{k|k}^J(\mathbf{L}_k)\} \text{ s.t. } \mathbf{L}_k \in \mathcal{L}_k, \tag{21}$$

and is computed according to (13a)-(13c), relying in part on the knowledge of

$$\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} = \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-2}}\mathbf{F}_{k-2}^H + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{w}_{k-2}} \tag{22a}$$

$$\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k} = \mathbf{F}_{k-1}\mathbf{C}_{\mathbf{x}_{k-1},\mathbf{v}_k} + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}. \tag{22b}$$

We can distinguish two possible cases (#1 and #2):

#1) $d\mathbf{F}_{k-1} \neq \mathbf{0}$, that is, $\mathbf{F}_{k-1} \neq \widehat{\mathbf{F}}_{k-1}$. In that situation, we have only access to the knowledge of

$$\begin{cases} \widehat{\mathbf{C}}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} = \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-2}}\widehat{\mathbf{F}}_{k-2}^H + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{w}_{k-2}} \\ \widehat{\mathbf{C}}_{\mathbf{x}_k,\mathbf{v}_k} = \widehat{\mathbf{F}}_{k-1}\mathbf{C}_{\mathbf{x}_{k-1},\mathbf{v}_k} + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k} \end{cases}.$$

Then, we must restrict to the "standard" LDSS model which satisfies (8), and we have that $\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} = \mathbf{0}$ and $\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k} = \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}$, leading to $\widehat{\mathbf{C}}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} = \mathbf{0}$ and $\widehat{\mathbf{C}}_{\mathbf{x}_k,\mathbf{v}_k} = \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}$. In this case, $\mathbf{P}_{k|k}^J(\mathbf{L}_k)$ (6) reduces to

$$\begin{aligned}\mathbf{P}_{k|k}^J(\mathbf{L}_k) &= (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)^H \\ &+ \mathbf{L}_k\mathbf{C}_{\mathbf{v}_k}\mathbf{L}_k^H - (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}\mathbf{L}_k^H \\ &- \mathbf{L}_k\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}^H(\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)^H, \end{aligned}\tag{23a}$$

$$\mathbf{P}_{k|k-1} = \widehat{\mathbf{F}}_{k-1}\mathbf{P}_{k-1|k-1}^b\widehat{\mathbf{F}}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}}, \tag{23b}$$

$$\mathbf{P}_{k-1|k-1}^b = \mathbb{E}[(\widehat{\mathbf{x}}_{k-1|k-1}^b - \mathbf{x}_{k-1})(\cdot)^H], \tag{23c}$$

and the solution of (21) only depends on $\widehat{\mathbf{F}}_{k-1}$, $\widehat{\mathbf{H}}_k$, $\mathbf{C}_{\mathbf{w}_k}$, $\mathbf{C}_{\mathbf{v}_k}$, $\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{v}_k}$, $\mathbf{C}_{\mathbf{x}_0}$ and $\mathbf{m}_{\mathbf{x}_0}$.

#2) $d\mathbf{F}_{k-1} = \mathbf{0}$, that is, $\mathbf{F}_{k-1} = \widehat{\mathbf{F}}_{k-1}$. In that situation, we have access to the knowledge of $\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}$ and $\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k}$, which means that we can use the most general form of $\mathbf{P}_{k|k}^J(\mathbf{L}_k)$ (6) derived under (4) and that the solution of (21) is given by (13a)-(13c).

To conclude, we analyze what happens at time $k = 1$. Again, we can distinguish two cases (#3 and #4):

#3) If $\mathbf{m}_{\mathbf{x}_0}$ and $\mathbf{C}_{\mathbf{x}_0}$ are (perfectly) known, then (17) is used, leading to (18a)-(18b) written as:

$$\begin{aligned}\widehat{\mathbf{x}}_{1|1}(\mathbf{L}_1) - \mathbf{x}_1 &= (\mathbf{I} - \mathbf{L}_1\widehat{\mathbf{H}}_1)(\widehat{\mathbf{F}}_0(\mathbf{m}_{\mathbf{x}_0} - \mathbf{x}_0) - \widehat{\mathbf{u}}_0 \\ &- \mathbf{w}_0) + \mathbf{L}_1\widehat{\mathbf{c}}_1 + \mathbf{L}_1\mathbf{v}_1 + \boldsymbol{\varepsilon}_1(\mathbf{L}_1) \end{aligned}\tag{24}$$

$$\begin{aligned}\boldsymbol{\varepsilon}_1(\mathbf{L}_1) &= \mathbf{L}_1(\widehat{\mathbf{H}}_1\boldsymbol{\eta}_0 + \mathbf{j}_1 + d\mathbf{c}_1) + \mathbf{L}_1 d\mathbf{H}_1(\mathbf{x}_1 + \boldsymbol{\eta}_0) \\ &- (\mathbf{I} - \mathbf{L}_1\widehat{\mathbf{H}}_1)(d\mathbf{F}_0\mathbf{x}_0 + d\mathbf{u}_0) \end{aligned}\tag{25}$$

#4) If $\mathbf{m}_{\mathbf{x}_0}$ and/or $\mathbf{C}_{\mathbf{x}_0}$ are (partially) unknown, then a Fisher initialization is used thus leading to:

$$\widehat{\mathbf{x}}_{1|1}(\mathbf{L}_1) - \mathbf{x}_1 = \mathbf{L}_1\mathbf{v}_1 + \boldsymbol{\varepsilon}_1(\mathbf{L}_1) \text{ s.t. } \mathbf{L}_1\widehat{\mathbf{H}}_1 = \mathbf{I}$$

$$\boldsymbol{\varepsilon}_1(\mathbf{L}_1) = \boldsymbol{\eta}_0 + \mathbf{L}_1 d\mathbf{H}_1(\mathbf{x}_1 + \boldsymbol{\eta}_0) + \mathbf{L}_1(\mathbf{j}_1 + d\mathbf{c}_1). \tag{26a}$$

In this case, $\boldsymbol{\eta}_0 \neq \mathbf{0} \Rightarrow \boldsymbol{\varepsilon}_1(\mathbf{L}_1) = \varnothing$. Therefore a necessary condition to allow $\boldsymbol{\varepsilon}_1(\mathbf{L}_1) \neq \varnothing$ is $\boldsymbol{\eta}_0 = \mathbf{0}$, leading to

$$\boldsymbol{\varepsilon}_1(\mathbf{L}_1) = \mathbf{L}_1 d\mathbf{H}_1\mathbf{x}_1 + \mathbf{L}_1(\mathbf{j}_1 + d\mathbf{c}_1). \tag{26b}$$

Finally, it is important to notice that if: 1a) $\mathbf{m}_{\mathbf{x}_0}$ and $\mathbf{C}_{\mathbf{x}_0}$ are (perfectly) known and $\mathcal{L}_1$, where $\boldsymbol{\varepsilon}_1(\mathbf{L}_1)$ is given by (25), is non empty; or 1b) $\mathbf{m}_{\mathbf{x}_0}$ and/or $\mathbf{C}_{\mathbf{x}_0}$ are (partially) unknown, $\boldsymbol{\eta}_0 = \mathbf{0}$, and $\mathcal{L}_1$, where $\boldsymbol{\varepsilon}_1(\mathbf{L}_1)$ is given by (26b), is non empty; and 2) $\mathcal{L}_k$, where $\boldsymbol{\varepsilon}_k(\mathbf{L}_k)$ is given by (18b), is non empty at each time $k \geq 2$; then *LCKF computed from the assumed LDSS model (1b) is matched to the true observations* $\mathbf{y}_k$ *(1c), and the recursion (13a)-(13c) minimizes the MSE associated with the nominal state* $\mathbf{x}_k$ *(1a). We then obtain the performance of LCKF for the assumed LDSS model (1b) with an increase of the achievable MSE due to the introduction of additional LCs* (19). The problem still remaining is how to define the set of constraints, associated to an appropriate non empty subset of gain matrices, to allow the mitigation of errors induced by the model mismatch, which is detailed in the sequel for several special cases of interest in many applications.

### B. Mitigation in LDSS Models where {True = Nominal}

We first consider the case where (1c)=(1a), that is {True = Nominal}. In other words, there are no unwanted disturbances into the system model ($\boldsymbol{\eta}_{k-1} = \mathbf{0}, \mathbf{j}_k = \mathbf{0}, \mathbf{x}_k'' = \mathbf{x}_k$), then the mismatch is only on the partially unknown system matrices and input values, and (18b) reduces to

$$\begin{aligned}\boldsymbol{\varepsilon}_k(\mathbf{L}_k) &= \mathbf{L}_k d\mathbf{H}_k\mathbf{x}_k + \mathbf{L}_k d\mathbf{c}_k \\ &- (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)(d\mathbf{F}_{k-1}\mathbf{x}_{k-1} + d\mathbf{u}_{k-1}). \end{aligned}\tag{27a}$$

In that case, a first constraint $\mathbf{L}_k d\mathbf{H}_k = \mathbf{0}$ leads to

$$\boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{L}_k d\mathbf{c}_k - (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)(d\mathbf{F}_{k-1}\mathbf{x}_{k-1} + d\mathbf{u}_{k-1})$$

then the complete set of constraints to have $\boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0}$ is

$$\begin{aligned}\mathbf{L}_k[d\mathbf{H}_k\ d\mathbf{c}_k] &= \mathbf{L}_k d\mathbb{H}_k = \mathbf{0} \\ (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)[d\mathbf{F}_{k-1}\ d\mathbf{u}_{k-1}] &= (\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)d\mathbb{F}_{k-1} = \mathbf{0}\end{aligned}\tag{27b}$$

There are two possible cases:

- Case 1) $rank\,(d\mathbb{F}_{k-1}) = P_k$

In this case, we have in (27b) that,

$$\{(\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)d\mathbb{F}_{k-1} = \mathbf{0},\ rank\,(d\mathbb{F}_{k-1}) = P_k\}$$

which implies that $\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k = \mathbf{0}$, then leading to a degenerated form of the LCKF recursion (17)

$$\widehat{\mathbf{x}}_{k|k}(\mathbf{L}_k) = \mathbf{L}_k(\mathbf{y}_k - \widehat{\mathbf{c}}_k), \tag{28a}$$

$$\widehat{\mathbf{x}}_{k|k}^b = \mathbf{L}_k^b(\mathbf{y}_k - \widehat{\mathbf{c}}_k), \tag{28b}$$

$$\mathbf{L}_k^b = \arg\min_{\mathbf{L}_k}\{\mathbf{L}_k\mathbf{C}_{\mathbf{v}_k}\mathbf{L}_k^H\} \text{ s.t. } \mathbf{L}_k[\widehat{\mathbf{H}}_k\ d\mathbb{H}_k] = [\mathbf{I}\ \mathbf{0}]. \tag{28c}$$

Thus if $rank\,(d\mathbb{F}_{k-1}) = P_k$, the introduction of LCs to mitigate modelling errors in state matrices $d\mathbb{F}_{k-1}$ removes the KF main merit, that is, the ability to combine previous observations to improve the estimation of the current state.

- Case 2) $rank\,(d\mathbb{F}_{k-1}) < P_k$

In this case, (27b) can be recast as $\{\mathbf{L}_k d\mathbb{H}_k = \mathbf{0},\ \mathbf{L}_k(\widehat{\mathbf{H}}_k d\mathbb{F}_{k-1}) = d\mathbb{F}_{k-1}\}$, and the LCKF recursion (17) does not degenerate as above. More specifically, let $d\mathbb{F}_{k-1} = \mathbb{U}_{k-1}d\boldsymbol{\Phi}_{k-1}$ be the SVD of $d\mathbb{F}_{k-1}$, where $\mathbf{U}_{k-1} \in \mathbb{C}^{P_k \times R_{k-1}}$ has full rank $R_{k-1} < P_k$ and $d\boldsymbol{\Phi}_{k-1} \in \mathbb{C}^{R_{k-1} \times P_k}$. Then

$$(\mathbf{I} - \mathbf{L}_k\widehat{\mathbf{H}}_k)\mathbb{U}_{k-1}d\boldsymbol{\Phi}_{k-1} = \mathbf{0},\ \forall d\boldsymbol{\Phi}_{k-1}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3132890, IEEE Transactions on Automatic Control



Since $\mathbb{U}_{k-1}$ has full rank the constraint above is equivalent to

$$(\mathbf{I} - \mathbf{L}_k \widehat{\mathbf{H}}_k)\mathbb{U}_{k-1} = \mathbf{0} \;\Leftrightarrow\; \mathbf{L}_k(\widehat{\mathbf{H}}_k \mathbb{U}_{k-1}) = \mathbb{U}_{k-1}$$

and (27b) becomes

$$\mathbf{L}_k[d\mathbb{H}_k \; \widehat{\mathbf{H}}_k \mathbb{U}_{k-1}] = [\mathbf{0} \; \mathbb{U}_{k-1}] \;\Rightarrow\; \boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0}. \quad (29)$$

Within the context of constrained filtering for the the case where (1c)=(1a), it is worth noting the following previous contributions: i) [25, (8)] mentions the LCs $(\mathbf{I} - \mathbf{L}_k \mathbf{H}_k)\, d\mathbf{u}_{k-1} = \mathbf{0}$ where $d\mathbf{u}_{k-1} = \mathbf{G}_{k-1}\mathbf{u}_{k-1}$ and $\mathbf{u}_{k-1}$ are unknown inputs; ii) considering $\mathbf{u}_{k-1}$ and $\mathbf{c}_k$ unknown inputs, the LCs

$$\begin{aligned} (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k)\, d\mathbf{u}_{k-1} = \mathbf{0}, \quad d\mathbf{u}_{k-1} = \mathbf{G}_{k-1}\mathbf{u}_{k-1}, \\ \mathbf{L}_k d\mathbf{c}_k = \mathbf{0}, \quad d\mathbf{c}_k = \boldsymbol{\Gamma}_k \mathbf{d}_k, \end{aligned}$$

are mentioned in [26], [27] and [9, Sec. III.E].

### C. Mitigation in LDSS Models where {Assumed = Nominal}

The second case of interest is when (1b)=(1a), that is {Assumed = Nominal}. In other words, there is no mismatch on the system matrices and input values ($d\mathbf{F}_{k-1} = \mathbf{0}, d\mathbf{H}_k = \mathbf{0}, d\mathbf{u}_{k-1} = \mathbf{0}, d\mathbf{c}_k = \mathbf{0}$), but we have unwanted disturbances in the true system model (1c), and (18b) reduces to

$$\boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{L}_k \mathbf{H}_k \mathbf{F}_{k-1}(\mathbf{x}''_{k-1} - \mathbf{x}_{k-1}) + \mathbf{L}_k(\mathbf{H}_k \boldsymbol{\eta}_{k-1} + \mathbf{j}_k) \quad (30a)$$

The set of constraints to mitigate the error is

$$\mathbf{L}_k \left[\mathbf{H}_k \mathbf{F}_{k-1} \; \mathbf{H}_k \boldsymbol{\eta}_{k-1} \; \mathbf{j}_k\right] = \mathbf{0} \;\Rightarrow\; \boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0}. \quad (30b)$$

### D. Generalization of the {True = Nominal} Case: Mitigation in LDSS Models where only $\boldsymbol{\eta}_{k-1} = \mathbf{0}$

In some cases it may be interesting to consider that we have a mismatch on system matrices and input values, and unwanted disturbances appear only on the observation equation (i.e., jammers), $\boldsymbol{\eta}_{k-1} = \mathbf{0}, \mathbf{x}''_k = \mathbf{x}_k$, and (18b) reduces to

$$\begin{aligned} \boldsymbol{\varepsilon}_k(\mathbf{L}_k) &= \mathbf{L}_k\left(d\mathbf{H}_k \mathbf{x}_k + \mathbf{j}_k + d\mathbf{c}_k\right) \\ &\quad - (\mathbf{I} - \mathbf{L}_k \widehat{\mathbf{H}}_k)\left(d\mathbf{F}_{k-1}\mathbf{x}_{k-1} + d\mathbf{u}_{k-1}\right), \end{aligned} \quad (31a)$$

which is of the same form as (27a). By analogy, the set of constraints to mitigate the error is given by

$$\left.\begin{aligned} \mathbf{L}_k\left[d\mathbb{H}_k \; \mathbf{j}_k\right] = \mathbf{0} \\ (\mathbf{I} - \mathbf{L}_k \widehat{\mathbf{H}}_k) d\mathbb{F}_{k-1} = \mathbf{0} \end{aligned}\right\} \;\Rightarrow\; \boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0}, \quad (31b)$$

which yields updated versions of (28c) or (29) according to $rank\,(d\mathbb{F}_{k-1})$. If $rank\,(d\mathbb{F}_{k-1}) < P_k$, then (31b) becomes

$$\mathbf{L}_k[d\mathbb{H}_k \; \mathbf{j}_k \; \widehat{\mathbf{H}}_k \mathbb{U}_{k-1}] = [\mathbf{0} \; \mathbf{0} \; \mathbb{U}_{k-1}] \;\Rightarrow\; \boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0}. \quad (31c)$$

### E. Generalization of the {Assumed = Nominal} Case: Mitigation in LDSS Models where only $d\mathbf{F}_{k-1} = \mathbf{0}$

It may be a reasonable assumption of interest to consider a mismatched model where only the process equation is correct, that is, $d\mathbf{F}_{k-1} = \mathbf{0}$. In this case, (18b) reduces to

$$\begin{aligned} &\boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{L}_k d\mathbf{H}_k \mathbf{x}''_k - (\mathbf{I} - \mathbf{L}_k \widehat{\mathbf{H}}_k) d\mathbf{u}_{k-1} + \\ &\mathbf{L}_k(\widehat{\mathbf{H}}_k \mathbf{F}_{k-1}(\mathbf{x}''_{k-1} - \mathbf{x}_{k-1}) + \widehat{\mathbf{H}}_k \boldsymbol{\eta}_{k-1} + \mathbf{j}_k + d\mathbf{c}_k), \end{aligned} \quad (32a)$$

which is of the same form as (27a) and (31a). Again, by analogy, the set of mitigation constraints is

$$\left.\begin{aligned} \mathbf{L}_k[d\mathbb{H}_k \; \widehat{\mathbf{H}}_k \boldsymbol{\eta}_{k-1} \; \mathbf{j}_k \; \widehat{\mathbf{H}}_k \mathbf{F}_{k-1}] = \mathbf{0} \\ (\mathbf{I} - \mathbf{L}_k \widehat{\mathbf{H}}_k) d\mathbf{u}_{k-1} = \mathbf{0} \end{aligned}\right\} \Rightarrow \boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0} \quad (32b)$$

which yields, by replacing $d\mathbb{F}_{k-1}$ by $d\mathbf{u}_{k-1}$, updated versions of (28c) or (29) according to $rank\,(d\mathbf{u}_{k-1})$.

### F. Mitigation of System Disturbances or Noises Uncertainty

For the sake of legibility, in the previous sections possible LCs to mitigate unwanted disturbances were summarized as

$$\mathbf{L}_k \mathbf{j}_k = \mathbf{0}, \; \mathbf{L}_k \mathbf{H}_k \boldsymbol{\eta}_{k-1} = \mathbf{0}, \; \mathbf{L}_k \widehat{\mathbf{H}}_k \boldsymbol{\eta}_{k-1} = \mathbf{0}. \quad (33a)$$

More precisely, LCs can mitigate unwanted and unknown disturbances if they lie in a known vector subspace. Indeed, if $\mathbf{j}_k = \boldsymbol{\Psi}_k \mathbf{i}_k$, $\boldsymbol{\Psi}_k$ known and $\mathbf{i}_k$ unknown, and/or $\boldsymbol{\eta}_{k-1} = \boldsymbol{\Phi}_{k-1}\mathbf{g}_{k-1}$, $\boldsymbol{\Phi}_{k-1}$ known and $\mathbf{g}_{k-1}$ unknown, then

$$\begin{aligned} \mathbf{L}_k \boldsymbol{\Psi}_k = \mathbf{0} \;&\Rightarrow\; \mathbf{L}_k \mathbf{j}_k = \mathbf{0}, \\ \mathbf{L}_k \mathbf{H}_k \boldsymbol{\Phi}_{k-1} = \mathbf{0} \;&\Rightarrow\; \mathbf{L}_k \mathbf{H}_k \boldsymbol{\eta}_{k-1} = \mathbf{0}, \\ \mathbf{L}_k \widehat{\mathbf{H}}_k \boldsymbol{\Phi}_{k-1} = \mathbf{0} \;&\Rightarrow\; \mathbf{L}_k \widehat{\mathbf{H}}_k \boldsymbol{\eta}_{k-1} = \mathbf{0}. \end{aligned} \quad (33b)$$

Interestingly enough, $\boldsymbol{\eta}_{k-1}$ and $\mathbf{j}_k$ may also account for system noises uncertainty where the uncertainties lie in known vector subspaces. Indeed, if we consider for instance the measurement noise, $(\mathbf{v}_k + \mathbf{j}_k)$ models a misspecification of the nominal/assumed measurement covariance matrix in comparison with the true one, i.e., a misspecification of the form

$$\mathbf{C}_{\mathbf{v}_k + \mathbf{j}_k} = \mathbf{C}_{\mathbf{v}_k} + \boldsymbol{\Psi}_k \mathbf{C}_{\mathbf{j}_k} \boldsymbol{\Psi}_k^H + \mathbf{C}_{\mathbf{v}_k, \mathbf{j}_k} \boldsymbol{\Psi}_k^H + \boldsymbol{\Psi}_k \mathbf{C}_{\mathbf{v}_k, \mathbf{j}_k}^H,$$

with $\boldsymbol{\Psi}_k$ known and $\{\mathbf{C}_{\mathbf{j}_k}, \mathbf{C}_{\mathbf{v}_k, \mathbf{j}_k}\}$ unknown. In this situation LCs as in (33b) mitigate the measurement noise misspecification since $\mathbf{L}_k \boldsymbol{\Psi}_k = \mathbf{0} \Rightarrow \mathbf{C}_{\mathbf{L}_k(\mathbf{v}_k + \mathbf{j}_k)} = \mathbf{C}_{\mathbf{L}_k \mathbf{v}_k}$. A similar rationale can be brought to light for the state noise.

### G. Unbiased (and Biased) Minimum Variance Estimation for Misspecified LDSS Systems

It is important to remark that the previous mitigation strategies, i.e., using LCs to cancel the error term $\boldsymbol{\varepsilon}_k(\mathbf{L}_k)$, can also be used to obtain an unbiased minimum variance estimation filter for misspecified LDSS models. Indeed, notice that

$$\boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0} \Rightarrow \mathbb{E}\left[\boldsymbol{\varepsilon}_k(\mathbf{L}_k)\right] = \mathbf{0}. \quad (34)$$

Or simply, one can directly resort to the set of LCs for which $\mathbb{E}\left[\boldsymbol{\varepsilon}_k(\mathbf{L}_k)\right] = \mathbf{0}$. In that case,

$$\begin{aligned} &\mathbb{E}\left[\boldsymbol{\varepsilon}_k(\mathbf{L}_k)\right] = \mathbf{L}_k \widehat{\mathbf{H}}_k(\widehat{\mathbf{F}}_{k-1} + d\mathbf{F}_{k-1})(\mathbf{m}_{\mathbf{x}''_{k-1}} - \mathbf{m}_{\mathbf{x}_{k-1}}) \\ &- (\mathbf{I} - \mathbf{L}_k \widehat{\mathbf{H}}_k)\left(d\mathbf{F}_{k-1}\mathbf{m}_{\mathbf{x}_{k-1}} + d\mathbf{u}_{k-1}\right) \\ &+ \mathbf{L}_k d\mathbf{H}_k\left(\widehat{\mathbf{u}}_{k-1} + d\mathbf{u}_{k-1}\right) \\ &+ \mathbf{L}_k d\mathbf{H}_k((\widehat{\mathbf{F}}_{k-1} + d\mathbf{F}_{k-1})\mathbf{m}_{\mathbf{x}''_{k-1}} + \mathbb{E}\left[\boldsymbol{\eta}_{k-1}\right]) \\ &+ \mathbf{L}_k(\widehat{\mathbf{H}}_k \boldsymbol{\eta}_{k-1} + \mathbb{E}\left[\mathbf{j}_k\right] + d\mathbf{c}_k), \end{aligned} \quad (35a)$$

and the optimal gain matrix corresponds to

$$\mathbf{L}_k^b = \arg\min_{\mathbf{L}_k}\{\mathbf{P}_{k|k}^J(\mathbf{L}_k)\} \text{ s.t. } \mathbb{E}\left[\boldsymbol{\varepsilon}_k(\mathbf{L}_k)\right] = \mathbf{0}, \quad (35b)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3132890, IEEE Transactions on Automatic Control



where the solution is obtained from $\mathbf{P}^J_{k|k}(\mathbf{L}_k)$ computed with the assumed LDSS model (1b). This implies that the proposed new LCKF approach allows to obtain an unbiased filter w.r.t. the nominal state, for which the MSE performances are the ones obtained for LCKF considering the assumed LDSS model (1b). In other words, this allows to obtain an unbiased filter under model mismatch. The main effect of LCs is then to recenter the assumed state, $\mathbf{x}'_k$, into the nominal one, $\mathbf{x}_k$.

In conclusion, we can also interpret that using (35b) one obtains an optimal biased minimum-variance estimation framework for misspecified LDSS systems, that is, from the assumed LDSS model (1b) this leads to a minimum MSE filter centred to the nominal state, $\mathbf{x}_k$.

## VI. Illustrative Example: Robust Navigation

In order to show the validity of the proposed robust LCKF we assess its performance in a navigation problem where the system is affected by a state input mismatch, $d\mathbf{u}_k \neq \mathbf{0}$, and a mismatched system model calibration, $d\mathbf{H}_k \neq \mathbf{0}$.

The benchmark example in [28], [9, VI-A] is explored, where the 2D position and velocity of a vehicle is estimated, $\mathbf{x}_k = (p_{x,k}, p_{y,k}, v_{x,k}, v_{y,k})^\top$, and the state is controlled by an acceleration command input $u_k$,

$$\mathbf{x}_{k+1} = \underbrace{\begin{pmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\mathbf{F}} \mathbf{x}_k + \underbrace{\begin{pmatrix} 0 \\ 0 \\ T\sin(\theta) \\ T\cos(\theta) \end{pmatrix}}_{\mathbf{G}(\theta)} u_k + \mathbf{w}_k$$

with $T = 1$ s, $\theta$ the vehicle heading angle and $\mathbf{w}_k$ a zero-mean Gaussian noise with covariance $\mathbf{Q}$. The vehicle is following a specific North-East direction (i.e., heading angle $\theta$), thus the state noise is constrained to this direction ($\sigma_p^2 = 1$, $\sigma_v^2 = 0.1$),

$$\mathbf{Q} = \begin{pmatrix} \sigma_p^2 \tan^2(\theta) & \sigma_p^2 \tan(\theta) & 0 & 0 \\ \sigma_p^2 \tan(\theta) & \sigma_p^2 & 0 & 0 \\ 0 & 0 & \sigma_v^2 \tan^2(\theta) & \sigma_v^2 \tan(\theta) \\ 0 & 0 & \sigma_v^2 \tan(\theta) & \sigma_v^2 \end{pmatrix}.$$

The system has access to two (potentially not well calibrated) sensors which provide position and velocity estimates on a different reference frame, then relative to the vehicle orientation $\psi$. The measurement equation is ($\boldsymbol{\psi} = (\psi_1, \psi_2)^\top$)

$$\mathbf{y}_k = \mathbf{H}(\boldsymbol{\psi})\mathbf{x}_k + \mathbf{n}_k, \; \mathbf{H}(\boldsymbol{\psi}) = \begin{bmatrix} \mathbf{H}_1(\psi_1) \\ \mathbf{H}_2(\psi_2) \end{bmatrix},$$

$$\mathbf{H}_i(\psi_i) = \begin{bmatrix} \cos(\psi_i) & \sin(\psi_i) & 0 & 0 \\ -\sin(\psi_i) & \cos(\psi_i) & 0 & 0 \\ 0 & 0 & \cos(\psi_i) & \sin(\psi_i) \\ 0 & 0 & -\sin(\psi_i) & \cos(\psi_i) \end{bmatrix},$$

and $\mathbf{n}_k$ a zero-mean Gaussian noise with covariance $\mathbf{R} = \mathrm{diag}(10, 10, 1, 1, 1, 1, 0.1, 0.1)$. The initial state is $\mathbf{x}_0 = (0, 0, 10\sin(\theta), 10\cos(\theta))^\top$ (i.e., velocity = 36 km/h in the heading direction), and the filters are initialized at $\widehat{\mathbf{x}}_{0|0} = (100, 100, 15, 15)^\top$ and $\mathbf{P}_{0|0} = \mathrm{diag}(50, 50, 5, 5)$.

In [9], [28] the heading angle $\theta$ was assumed to be perfectly known, but in practice we may have a real trajectory with a deviation from the nominal (i.e., wrong heading angle estimation), then $\widehat{\theta} = \theta + d\widehat{\theta}$, which corresponds to an input mismatch. In addition, we consider a sensor platform calibration error on the second sensor measurement reference system, $\widehat{\psi}_2 = \psi_2 + d\widehat{\psi}_2$ (i.e., $\widehat{\boldsymbol{\psi}} = (\psi_1, \widehat{\psi}_2)^\top$). Then we have the following Nominal-Assumed LDSS pair:

$$\text{Nominal}: \begin{cases} \mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}(\theta)u_k + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{H}(\boldsymbol{\psi})\mathbf{x}_k + \mathbf{n}_k \end{cases}$$

$$\text{Assumed}: \begin{cases} \mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}(\widehat{\theta})u_k + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{H}(\widehat{\boldsymbol{\psi}})\mathbf{x}_k + \mathbf{n}_k \end{cases}$$

Both $\mathbf{G}(\theta)$ and $\mathbf{H}(\boldsymbol{\psi})$ can be approximated as

$$\mathbf{G}(\theta) \simeq \begin{pmatrix} 0 \\ 0 \\ T\sin(\widehat{\theta}) \\ T\cos(\widehat{\theta}) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ T\cos(\widehat{\theta}) \\ -T\sin(\widehat{\theta}) \end{pmatrix} d\widehat{\theta},$$

$$\mathbf{H}_i(\psi_i) \simeq \mathbf{H}_i(\widehat{\psi}_i) - \begin{bmatrix} \mathbf{D}(\widehat{\psi}_i)^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{D}(\widehat{\psi}_i)^\top \end{bmatrix} d\widehat{\psi}_i,$$

with $\mathbf{D}(\psi_i) = \begin{bmatrix} -\sin(\psi_i) & -\cos(\psi_i) \\ \cos(\psi_i) & -\sin(\psi_i) \end{bmatrix}$. If we define $\mathbf{r} = (0, 0, T\cos(\widehat{\theta}), -T\sin(\widehat{\theta}))^\top$, then $\widehat{\mathbf{u}}_k = \mathbf{r}u_k$ and $d\mathbf{u}_k = \mathbf{r}u_k d\widehat{\theta}$. Because $rank(d\mathbf{u}_k) < P_k$ the input part of the constraint (29) is $\mathbf{L}_k\mathbf{H}_k(\widehat{\boldsymbol{\psi}})\mathbf{r} = \mathbf{r}$. In order to mitigate the measurement model mismatch, the LCs (29) is given by $\mathbf{L}_k d\mathbf{H}_k = \mathbf{0}$, which in this case simply deletes the mismatched measurement from the LCKF estimate, which is not what we are looking for. Because this is a {True = Nominal} case (i.e., $\boldsymbol{\eta}_k = \mathbf{0}$, $\mathbf{j}_k = \mathbf{0}$), instead of looking for the LCs for which $\boldsymbol{\varepsilon}_k(\mathbf{L}_k) = \mathbf{0}$, we can enforce that $\mathbb{E}[\boldsymbol{\varepsilon}_k(\mathbf{L}_k)] = \mathbf{0}$ (refer to Section V-G for unbiased minimum variance estimation), which leads to $\mathbf{L}_k d\mathbf{H}_k \mathbf{m}_{\mathbf{x}_k} = \mathbf{0}$, and therefore the complete set of LCs for the robust LCKF is $\mathbf{L}_k\boldsymbol{\Delta}_k = \mathbf{T}_k$ with

$$\boldsymbol{\Delta}_k = \left[ \mathbf{H}(\widehat{\boldsymbol{\psi}})\mathbf{r} \; \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{D}^\top(\widehat{\psi}_2) & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^\top(\widehat{\psi}_2) \end{bmatrix} \mathbf{m}_{\mathbf{x}_k} \right], \mathbf{T}_k = [\mathbf{r} \; \mathbf{0}].$$

Results for a 5 minutes trajectory are obtained with the following setup: i) true heading angle $\theta = 60^\circ$ and true platform orientation $\psi_1, \psi_2 \in [-\pi, \pi]$, ii) alternate acceleration input $u_k$ changing from $+1$ m/s$^2$ to $-1$ m/s$^2$ every 30 seconds, and iii) $d\widehat{\theta} \in [-2\pi/10, 2\pi/10]$, max. error of $36^\circ$ w.r.t. $\theta$, and $d\widehat{\psi}_2 = 3.6^\circ$ w.r.t. $\psi_2$. We compare the following algorithms: 1) optimal KF with true knowledge of the system $\theta$ and $\psi_2$, 2) mismatched KF which considers $\widehat{\theta}$ and $\widehat{\psi}_2$, 3) a LCKF which only mitigates the impact of $\widehat{\theta}$ (LCKF1), 4) a LCKF which only mitigates the impact of $\widehat{\psi}_2$ (LCKF2), and 5) a LCKF which mitigates the impact of both $\widehat{\theta}$ and $\widehat{\psi}_2$ (LCKF3).

A trajectory example and the corresponding estimates are shown in Figure 1 (top). The 2D total mean RMSE over the trajectory for the different methods is: 1) RMSE optimal KF = 0.3, 2) RMSE mismatched KF = 10.36, 3) RMSE LCKF1 = 9.64, 4) RMSE LCKF2 = 1.87, and 5) RMSE LCKF3 = 0.33.

We can see that only mitigating the mismatched input (LCKF1), the filter is able to mitigate the fluctuations induced by the wrong acceleration but not the misalignment with

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3132890, IEEE Transactions on Automatic Control



the true coordinate frame. In contrast, mitigating the miscalibration error (LCKF2), the filter preserves the appropriate coordinate frame but does not correctly deal with the acceleration fluctuations induced by the input mismatch. Exploiting both constraints (LCKF3) the filter is able to mitigate both input and calibration mismatches, which confirms the LCKF performance improvement in mismatched LDSS models. To complete the discussion, the 2D RMSE w.r.t. the trajectory time is shown in Figure 1 (bottom). This result confirms the remarkable impact that a model mismatch may have on the final mismatched KF performance, and the interest of using appropriate LCs within the new LCKF framework.

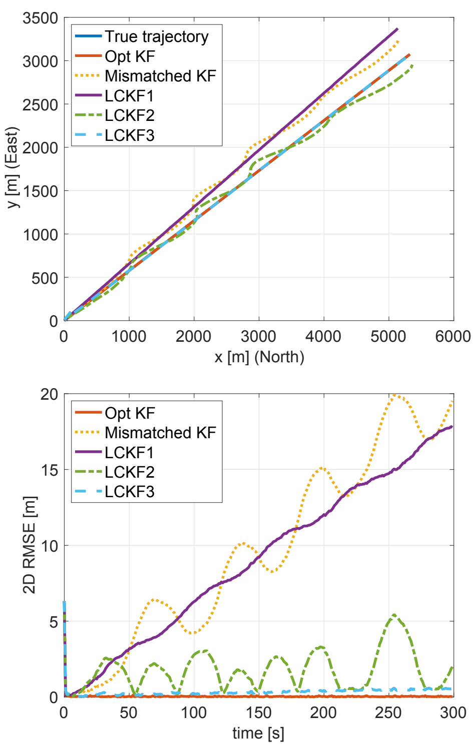


Fig. 1. (Top) Example of horizontal (x,y) trajectory and the corresponding estimates. (Bottom) 2D RMSE w.r.t. trajectory time using both optimal KF (no mismatch), mismatched KF and different LCKFs.

## VII. Conclusion

It is well-know that KF is very sensitive to model mismatch, that is, misspecified covariances, uncertainties in system matrices, filter initialization, uncertain inputs or unwanted system behaviors. In this contribution we explored the use of LCs for KF robustification. First, we introduced a general class of LCKFs, which allows at every time step to include or not non-stationary constraints. These LCKFs generalize previous results on information filtering, MVDR, GCKF, and can also be used to obtain unbiased minimum-variance estimators. In contrast to previous contributions, the goal was not to constrain the state or measurements, but rather use LCs to robustify KF under model mismatch. For that purpose, the impact of possible model mismatch was shown and how to use the LCKF in order to mitigate such mismatch was thoroughly discussed. Notice that these results may be of broad interest given that are derived from the most general KF form. Indeed, because such LCs can be taken into account in any KF generalization, this work sets the basis for a new robust filtering framework.

## References

[1] H. L. Van Trees, *Optimum Array Processing*, Wiley-Interscience, New York, USA, 2002.
[2] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Springer, Switzerland, 4th edition, 2013.
[3] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*, Wiley InterScience, New York, USA, 2006.
[4] H. Heffes, "The effect of erroneous models on the Kalman filter response," *IEEE Trans. Automatic Control*, vol. 11, no. 13, 1966.
[5] Y. Theodor and U. Shaked, "Robust discrete-time minimum-variance filtering," *IEEE Trans. Signal Process.*, vol. 44, no. 2, 1996.
[6] E. Chaumette et al., "On LMVDR Estimators For LDSS Models: conditions for existence and further applications," *IEEE Trans. Automatic Control*, vol. 64, no. 6, pp. 2598–2605, 2019.
[7] E. Chaumette, F. Vincent, and J. Vilà-Valls, "Linearly Constrained Wiener Filter Estimates For Linear Discrete State-Space Models," in *Proc. IEEE Asilomar*, Nov. 2018.
[8] J. Vilà-Valls, D. Vivet, E. Chaumette, F. Vincent, and P. Closas, "Recursive Linearly Constrained Wiener Filter for Robust Multi-Channel Signal Processing," *Signal Processing*, vol. 167, Feb. 2020.
[9] B. O. S. Teixeira et al., "Gain-constrained Kalman filtering for linear and nonlinear systems," *IEEE Trans. Sig. Process.*, vol. 56, no. 9, 2008.
[10] D. Dardari, P. Closas, and P. M. Djuric, "Indoor Tracking: Theory, Methods, and Technologies," *IEEE Trans. on Vehicular Technology*, vol. 64, no. 4, pp. 1263–1278, April 2015.
[11] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312–2322, 2018.
[12] S. Haykin, *Kalman Filtering and Neural Networks*, vol. 47, John Wiley & Sons, 2004.
[13] H. Guo, D. Cao, H. Chen, C. Lv, H. Wang, and S. Yang, "Vehicle Dynamic State Estimation: State of the Art Schemes and Perspectives," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, 2018.
[14] J. Duník, S. K. Biswas, A. G. Dempster, T. Pany, and P. Closas, "State Estimation Methods in Navigation: Overview and Application," *IEEE Aerospace and Electronic Systems Magazine*, vol. 35, no. 12, 2020.
[15] L. Xie, L. Lu, D. Zhang, and H. Zhang, "Improved robust $H_2$ and $H_\infty$ filtering for uncertain discrete-time systems," *Automatica*, vol. 40, no. 5, pp. 873–880, 2004.
[16] R. Banavar, *A game theoretic approach to linear dynamic estimation*, Ph.D. thesis, University of Texas at Austin, 1992.
[17] X.-M. Shen and L. Deng, "Game theory approach to discrete $H_\infty$ filter design," *IEEE Trans. Signal Process.*, vol. 45, no. 4, 1997.
[18] Y. S. Shmaliy, "Linear optimal FIR estimation of discrete time-invariant state-space models," *IEEE Trans. Signal Process.*, vol. 58, no. 6, 2010.
[19] Y. S. Shmaliy, "An Iterative Kalman-Like Algorithm Ignoring Noise and Initial Conditions," *IEEE Trans. Sig. Process.*, vol. 59, no. 6, 2011.
[20] D. Simon and Y. S. Shmaliy, "Unified forms for Kalman and finite impulse response filtering and smoothing," *Automatica*, vol. 49, 2013.
[21] S. Zhao, Y. S. Shmaliy, B. Huang, and F. Liu, "Minimum variance unbiased FIR filter for discrete time-variant systems," *Automatica*, vol. 53, pp. 355–361, 2015.
[22] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust Statistics for Signal Processing*, Cambridge University Press, 2018.
[23] H. Fang et al., "Nonlinear Bayesian Estimation: from Kalman Filtering to a Broader Horizon," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 401–417, 2018.
[24] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, New Jersey, USA, 2000.
[25] P. K. Kitanidis, "Unbiased minimum-variance linear state estimation," *Automatica*, vol. 23, no. 6, pp. 775–778, 1987.
[26] M. Darouach, M. Zasadzinski, and M. Boutayeb, "Extension of minimum variance estimation for systems with unknown inputs," *Automatica*, vol. 39, no. 5, pp. 867–876, 2003.
[27] S. Gillijns and B. L. R. De Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough," *Automatica*, vol. 43, pp. 934–937, 2007.
[28] D. Simon and T. L. Chia, "Kalman filtering with state equality constraints," *IEEE Trans. Aerospace & Elec. Syst.*, vol. 38, no. 1, 2002.